Assessing the Efficacy of Improvements in User Satisfaction for Mobile Applications

User Feedback from the Review Data

Bomi Song, Woori Han and Yongtae Park

Department of Industrial Engineeing, Seoul National University, 1 Gwanka-ro, Gwanak-gu, Seoul, Republic of Korea

Keywords: Improvement of Mobile Application, User Satisfaction, User Feedback, Review Data.

Abstract: Monitoring user feedback is of central importance in the iterative and incremental appraoch to improvement of mobile applications. Knowledge from user review on mobile applications can be fruitful source of user feedback on firms' effort to improve their mobile applications. This paper proposes an approach to assessing the efficacy of improvements in user satisfaction for mobile applications using user review data. Specifically, overall satisfaction score and frequencies of updated features in review data are compared before and after updates of mobile applications. We believe our method can facilitate utilizing user reviews to track user feedback and obtain useful knowledge in planning and managing updates of mobile applications, and serve as a starting point of more general model.

1 INTRODUCTION

The strategic importance of achieving user satisfaction has become more apparent in mobile applications given the rapidly changing markets and fierce competition (Lal et al., 2001; Heyes, 2002). Chasing the growing markets for mobile applications and opportunities for new business, numerous mobile applications have appeared, but many of them have disappeared from the competition. Users are impatient with poor mobile applications due to low switching barrier and a lot of alternatives (Kimbler, 2010; Holzer and Ondrus, 2011).

In this situation, companies and mobile application developers are focusing increasing attention on iterative and incremental approach to development and deployment of a mobile application to develop innovative and useful services as well as deliver quality and reliable services (Abrahamsson et al., 2003; Abrahamsson, 2005; Rahimian and Ramsin, 2008). High flexibility of the iterative and incremental approach enables developers and providers of mobile applications to embrace environmental changes including technologies, customer needs, and regulations into mobile applications (Blazevic et al., 2003). Also, it provides a way for fast time-to-market of an innovation (Blazevic et al., 2003) and ongoing effort to innovate in a fast manner (Tapscott et al., 1988; Blazevic et al., 2003). To sum up, continual improvements by updating versions of mobile applications can increase loyalty of existing users, as well as attract new users with relative low cost than launching a complete service at the first time.

Monitoring user feedback is of central importance in the iterative and incremental appraoch to improvement of mobile applications. It is the first step for the next improvement, providing evidences for assessing the current improvements and establishing direction of the next improvement in the light of users (Highsmith and Cockburn, 2001; Lindvall et al., 2002). Also, as mentioned earlier, since they are rapidly changing, user requirements should be reassessed and reprioritized based on the user feedback (Bajic and Lyons, 2011). In this regard, understanding of how firms' improvement effort affects user satisfaction and mirroring the results on the next version are crucial in the success of mobile applications. However, as for the mobile applications, it is impeded in many cases due to a lack of quantitative data and systematic processes.

With respect to the data, there exist numerous methods for getting user feedback such as focus groups, large-scale test bed, and conference with users. However, small and medium enterprises in the

Song B., Han W. and Park Y..
Assessing the Efficacy of Improvements in User Satisfaction for Mobile Applications - User Feedback from the Review Data.
DOI: 10.5220/0003992105030506
In *Proceedings of the 9th International Conference on Informatics in Control, Automation and Robotics* (ICINCO-2012), pages 503-506
ISBN: 978-989-8565-22-8

mobile application industry, which are the majority, are hard to bear a large amount of cost of such direct relationships with users (Bajic and Lyons, 2011). In this situation, user reviews can provide for mobile applications developers and providers with useful information on users' thinking at relatively low cost. In addition, review data have several advantages as source of voice of customers high availability, validity, and influence (Bickart and Schindler, 2001; Büyüközkan et al., 2007).

This study primarily aims to propose a new approach to assess the efficacy of improvements in user satisfaction of mobile applications by analyzing user feedback in review data. Specifically, overall satisfaction score and frequencies of updated features in review data are compared before and after updates of mobile applications.

2 PROPOSED APPRAOCH

As shown in Figure 1, the proposed approach consists of four steps: *construction of database*, *development of keyword vectors*, *comparison before and after updates*, and *assessment of improvements*.

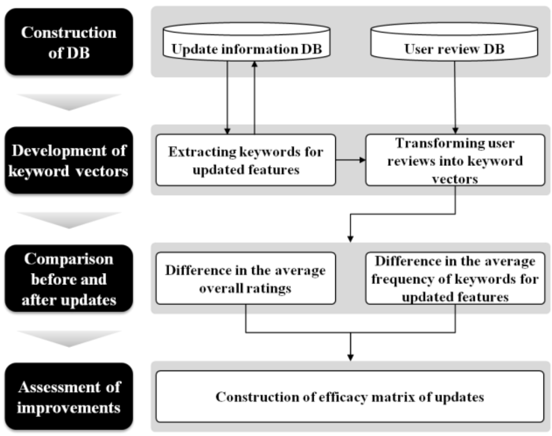

Figure 1: Overall process.

2.1 Step 1: Construction of Database

Two kinds of data are required: update information on mobile applications and user reviews on those mobile applications. On the on hand, update information can be found in a variety of sources on the Web such as official websites of application developers or providers, websites that provide information on mobile applications collectively (e.g. *appshopper.com*), and news sites dealing with mobile applications (e.g. www.cnet.com). On the other hand, basically, user reviews can be found in download sites of mobile applications like Apple App Store and Android Market. Blogs and forums can also be source of user reviews.

Such data are in the different formats according to their sources, as well as being in the unstructured textual format. Hence, update information and user reviews on the Web are stored in separate databases, named *updated information DB* and *User review DB*. Updated information DB consists of six fields including *application name*, *category*, *version*, *release date*, *price*, and *description of improvement*. User review DB consists of seven fields including *application name*, *version*, *reviewer*, *review date*, *overall rating*, *title*, and *comment*. To construct those databases, web documents including update information or user reviews on the target mobile applications are firstly collected. And then, the information that corresponds to each field of two databases is extracted from the documents by parsing them.

2.2 Step 2: Development of Keyword Vectors

In this step, user reviews are transformed into keyword vectors. The ultimate purpose of this study is to investigate how users think about the improvements achieved by the update of mobile applications that is represented in the review data. Therefore, in this study, keyword vectors for user reviews are designed to consist of *application name*, *version*, *overall rating*, and *frequency of keywords for updated features*.

While the first three fields are simply bring from the user review DB, the last filed needs further processes as follows. Firstly, keywords representing updated features should be identified in advance. For this purpose, keywords are extracted from the field of description of improvement of update information DB for each version of each application. Then, the

Table 1: An example of keyword vectors.

Application name	Version	Overall rating	Frequency of keywords for updated features			
			V1		V2	
			k1 (R)	k2 (R)	k3 (F)	k4 (R)
A1	V1	2	0	2	1	0
A1	V1	4	1	0	2	0
A1	V2	6	2	3	1	2
A1	V2	4	2	2	2	3
…	…	…	…	…	…	…

extracted keywords are annotated with types of them; new features (N), refinements (R), or fixing bugs/errors (F). Secondly, the frequency of keywords for updated features is counted from the field of title and comment of user review DB. Table 1 shows an example of keyword vectors.

2.3 Step 3: Comparison before and after Updates

Based on the developed keyword vectors, effect of a particular update on users' thinking is examined at the update level and feature level. To this end, average overall rating and average frequency of keywords for updated features are compared before and after updates, respectively.

For the update-level comparison, the hypothesis is; the average overall rating for a version of mobile application is different from that for its previous version. Since overall ratings can be considered a kind of Likert scale, Wilcoxon rank-sum test (also called Mann-Witney U test), a non-parametric statistical test, is used for testing the hypothesis. If the hypothesis is rejected, it is difficult to derive any conclusion about the efficacy of update without additional information. In contrast, if the hypothesis is accepted, it can be concluded that the update have affected the changes in average overall ratings, and further analysis is needed for identifying that which features improved by the update have influence in the change of overall rating. Hence, the feature-level comparison is needed for the version about which the update-level hypothesis is accepted.

The hypothesis for feature-level comparison is; the average frequency of keywords for updated features in the reviews for a version of mobile application is different from that for its previous version. Considering frequency of keyword ratio scale, independent samples Student's t-test is used for testing the hypothesis. The result might be one of four cases: the hypothesis is rejected and frequency of keywords remains high; the hypothesis is rejected and frequency of keywords remains low; the hypothesis is accepted and frequency of keywords is increased after the update; and the hypothesis is accepted and frequency of keywords is decreased after the update. The frequency of keywords for updated features appearing in user reviews per se, however, is difficult to provide information on whether users like the updated features or not since user feedback on the updated features can be positive or negative. Therefore, the change in frequency of keywords for updated features should be investigated in conjunction with the changes in overall rating.

2.4 Step 4: Assessment of Improvement

Based on the results of update-level and feature-level comparison before and after updates, *efficacy matrix of updates* is constructed. An example of efficacy matrix of updates is presented in Table 2.

Table 2: An example of efficacy matrix of updates.

		Change in frequency of keywords for updated features			
		Decrease	Increase	Remain high	Remain low
Change in overall rating after update	Increase	(a)	(c)	(e)	(g)
	Decrease	(b)	(d)	(f)	(h)

The row indicates direction of change in overall rating after update, which has been examined in the update-level comparison. On the other hand, the column represents change in frequency of keywords for updated features, which is obtained as the result of feature-level comparison. According to the results of comparison before and after comparisons in the previous step, each keyword for the updated feature is classified into one of eight cells in the efficacy matrix of updates.

On the efficacy matrix of updates, change in overall rating after update is regarded as the proxy of user satisfaction and change in frequency of keywords for updated features is regarded as the proxy of users' interest. By jointly investigating the changes in overall rating and keyword frequency, users' attitudes towards the updated features can be inferred for each cell of the efficacy matrix of updates as follows:

- Cell (a): users had often expressed their dissatisfaction with the features, but have rarely expressed their satisfaction with the features after improvement;
- Cell (b): users had often expressed their satisfaction with the features, but have rarely expressed their dissatisfaction with the features after improvement;
- Cell (c): users had not expected, but have been delighted with the updated features;
- Cell (d): Users have not liked 1) the updated features per se or 2) the quality of the updated features;
- Cell (e): complaints on the features have been addressed via the updates, and users has been satisfactory with the result of improvements;
- Cell (f): users had wanted the features to be improved, but have not liked the quality of the improvements;
- Cell (g) and (h): users might not be interested in the updated features, or the changes in

overall rating might be affected by the other improvements.

3 CONCLUSIONS

We have proposed a way of utilizing user reviews for assessing the efficacy of improvement in user satisfaction for mobile application. The changes in user satisfaction for updates and users' interests for the updated features were statistically examined by comparing the average overall satisfaction and the average frequency of keyword for the updated features, respectively. Moreover, the suggested efficacy matrix of updates enabled to assess users' attitudes towards the updated features. Our suggested approach can facilitate utilizing user reviews to track user feedback and obtain useful knowledge in planning and managing updates of mobile applications, or further in developing of new mobile applications.

The future research may include the following themes. Firstly, measuring user satisfaction on the individual updated features rather than overall satisfaction can allow more sophisticated analysis of efficacy matrix. For this purpose, sentiment analysis or conjoint analysis can be incorporated into the proposed approach. Secondly, indexes that represent the efficacy of updated features based on the efficacy matrix can provide more clear understanding of the updated features' influence on the user satisfaction. Finally, the suggested approach should be actually implemented with the appropriate cases. These topics can be fruitful area for future research.

ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No. 2009-0085757).

REFERENCES

Abrahamsson, P., 2005. Keynote: Mobile software development — The business opportunity of today. In *Proceedings of the International conference on Software Development*, Reykjavik, Iceland, 20-23.

Abrahamsson, P., Warsta J., Spionen, M. T., Ronkainen, J., 2003. New directions on agile methods: a comparative analysis. In *Proceedings of International Conference on Software Engineering (ICSE25)*, Portland, Oregon.

Bajic, D., Lyons, K., 2011. Leveraging social media to gather user feedback for software development. In *Proceedings of the 2nd International Workshop on Web 2.0 for Software Engineering (Web2SE 2011)*, Waikiki, Honolulu, HI, USA, 1-6.

Bickart, B., Schindler, R. M., 2001. Internet forums as influential sources of consumer information. *Journal of Interactive Marketing*, 15(3), 31-40.

Blazevic, V., Lievens, A., Klein, E., 2003. Antecedents of project learning and time-to-market during new mobile service development. *International Journal of Service Industry Management*, 14(1), 120–147.

Büyüközkan, G., Baykasoğlu, A., Dereli, T., 2007. Integration of internet and web-based tools in new product development process. *Production Planning and Control*, 1(1), 44-53.

Heyes, I. S., 2002. *Just Enough Wireless Computing*, Prentice Hall. Upper Saddle River.

Highsmith, J., Cockburn, A., 2001. Agile software development: the business of innovation. *IEEE Computer*, 34(9), 120-127.

Holzer, A., Ondrus, J., 2011. Mobile application market: A developer′s perspective. *Telematics and Informatics*, 28(1), 22-31.

Kimbler, K., 2010. App store strategies for service providers. In *Proceedings of the 14th International Conference on Intelligence in Next Generation Networks (ICIN 2010)*, Berlin, 1-5.

Lal, D., Pitt, D. C., and Beloucif, A., 2001. Restructuring in European telecommunications: modeling the evolving market. *European Business Review*, 13(3), 152-156.

Lindvall, M., Basili, V., Boehm, B., Costa, P., Dangle, K., Shull, F., Tesoriero, R., Williams, L., Zelkowitz, M., 2002. Empirical findings in agile methods. In *Proceedings of XP Universe and First Agile Universe Conference*, 197-207.

Rahimian, V., Ramsin, R., 2008. Designing an agile methodology for mobile software development: a hybrid method engineering approach. In *Proceedings of the 2nd International Conference on Research Challenges in Information Science (RCIS 2008)*, Marrakech, 337-342.

Tapscott, D., Lowy, A., Ticoll, D., 1998. *Blueprint to The Digital Economy: Creating Wealth in The Era of E-business*, McGraw-Hill. New York.